DOCUMENT RESUME

ED 069 693 TM 002 144

AUTHOR	Shafto, Michael
TITLE	Cluster Analysis by Linear Contrasts.
INSTITUTION	Educational Testing Service, Princeton, N.J.
SPONS AGENCY	National Science Foundation, Washington, D.C.
REPORT NO	ETS-RB-72-35
PUB DATE	Aug 72
NOTE	26p.; A draft
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Bibliographies; Classification; *Cluster Analysis; Cluster Grouping; *Correlation; Data Analysis; Factor Analysis; *Models; *Research Methodology; *Statistical Analysis; Tables (Data); Technical Reports
IDENTIFIERS	IICA; Interactive Intercolumnar Correlation Analysis

ABSTRACT

The purpose of this paper is to suggest a technique of cluster analysis which is similar in aim to the Interactive Intercolumnar Correlation Analysis (IICA), though different in detail. Two methods are proposed for extracting a single bipolar factor (a "contrast component") directly from the initial similarities matrix. The advantages of this general approach are that: (a) It helps avoid certain misclassification problems inherent in IICA; (b) It is related in a straightforward way to conventional techniques of multidimensional scaling and therefore allows a unified treatment of dimensional and "typal" structures; and (c) It provides an interesting solution to the problem of relations among linear contrasts based on different subsets of the stimuli. (Author/DJ)

RB-72-35

RESEARCH BULLETIN

CLUSTER ANALYSIS BY LINEAR CONTRASTS

Michael Shafto

Princeton University
and Educational Testing Service

This Bulletin is a draft for interoffice circulation. Corrections and suggestions for revision are solicited. The Bulletin should not be cited as a reference without the specific permission of the author. It is automatically superseded upon formal publication of the material.

Educational Testing Service
Princeton, New Jersey
August 1972

Errata for

CLUSTER ANALYSIS BY LINEAR CONTRASTS

(RB-72-35)

Michael Shafto

1. Page 9, equation (26)

now reads:

"$F(\underline{b}) = \underline{b}'R\underline{b} - [tr(R)/n]\ (\underline{b}'\underline{1})$"

should read:

"$F(\underline{b}) = \underline{b}'R\underline{b} - [tr(R)/n]\ (\underline{b}'\underline{1})^2$"

2. Page 9, paragraph immediately following equation (26'), second sentence,

now reads:

"This solution is invariant under transformations of the form $r_{ij} \to ar_{ij} + c$, while the Method A solution is invariant only under transformations of the form $r_{ij} \to ar_{ij}$."

should read:

"Solutions by either Method A or Method B are invariant under transformations of the form $r_{ij} \to ar_{ij} + c$."

November 17, 1972

CLUSTER ANALYSIS BY LINEAR CONTRASTS[1]

Michael Shafto

Princeton University and Educational Testing Service

Introduction

L. L. McQuitty (1967) has suggested a technique of hierarchical cluster analysis called Iterative Intercolumnar Correlational Analysis (IICA). McQuitty and Clark have provided a discussion of the mathematics of this technique, its application to real and artificial data, and its advantages and disadvantages in comparison with other methods (Clark & McQuitty, 1970; McQuitty, 1971; McQuitty, Abeles, & Clark, 1970; McQuitty & Clark, 1968). Coles and Stone (1972) have suggested a related technique.

IICA begins with a raw data matrix from which a first-order similarities matrix $R^{(1)}$ is computed (or, in some cases, $R^{(1)}$ may be obtained directly by subjects' similarity judgments). The larger the ij -th entry in $R^{(1)}$, the more "alike" or "similar," in some sense, stimuli i and j are judged to be. A second-order similarities matrix $R^{(2)}$ is then computed by intercorrelating the columns of $R^{(1)}$. That is, the ij -th entry in $R^{(2)}$ is the product-moment correlation between columns i and j of $R^{(1)}$. Then $R^{(3)}$ is computed by intercorrelating the columns of $R^{(2)}$, and so on, until a matrix $R^{(K)}$ is produced in which all elements have absolute value unity, within reasonable tolerance limits. The stimuli are then partitioned into two subsets: Those with +1 in the first column of $R^{(K)}$ go in one subset; those with -1 go in the other. (*Any* column could be used, not just the first.) A discussion of the convergence problem may be found in Clark and McQuitty (1970).

---

[1]Research supported by National Science Foundation Grant GB 8023X with Princeton University, project on "Mathematical Techniques in Psychology," Harold Gulliksen, Principal Investigator, and by Educational Testing Service.

The author is grateful to Dr. Walter Kristof for many valuable suggestions and comments on an earlier version of this report. The assistance of Mrs. Ann King in supervising the preparation of the manuscript, and of Mr. Terry Birch in drawing the figures, is also gratefully acknowledged.

The purpose of this paper is to suggest a technique of cluster analysis which is similar in aim to IICA, though different in detail. Two methods will be proposed for extracting a single bipolar factor (a "contrast component") directly from the initial similarities matrix. The advantages of this general approach are that (a) it helps avoid certain misclassification problems inherent in IICA; (b) it is related in a straightforward way to conventional techniques of multidimensional scaling (Torgerson, 1958) and therefore allows a unified treatment of dimensional and "typal" structures, and (c) it provides an interesting solution to the problem of relations among linear contrasts based on different subsets of the stimuli. This last problem was initially raised by McQuitty (1967).

I. Matrix Algebra of One IICA Iteration

The following discussion of the matrix algebra of one IICA cycle is intended to clarify the relationship between IICA and the new techniques outlined below.

Consider a typical iteration, starting with $R^{(k)}$ and ending with $R^{(k+1)}$. The superscript will be dropped from $R^{(k)}$ for purposes of the following discussion.

R is an $n \times n$ symmetric matrix which is positive semidefinite for $k \geq 2$. Therefore, there exist matrices U and D, such that D is diagonal, $U' = U^{-1}$, and $R = UDU'$.

Let $\underline{1}$ be a column vector of n 1's, and let I be the $n \times n$ identity matrix. Define

$$(1) \qquad M = I - \underline{1}\underline{1}'/n \quad ;$$

$$(2) \qquad \overline{R} = MR \quad ;$$

(3) $A = \bar{R}'\bar{R} = UDU'M'MUDU' = UDU'MUDU'$, since M is idempotent;

(4) $T = U'MU$; and

(5) $\underline{b} = U'\underline{1}$.

Thus,

(6) $T = I - \underline{bb}'/n$,

and the general element of T is

(7) $t_{ij} = \delta_{ij} - b_i b_j/n$,

where δ_{ij} is the Kronecker delta.

By the Cauchy-Schwarz inequality,

(8) $|b_i| \leq \sqrt{n}$, $i = 1,\ldots,n$.

Therefore,

(9) $|b_i b_j| \leq n$, $i,j = 1,\ldots,n$.

Now suppose that the inequality in (8) holds for each i , as it almost always will with real data. Then all the diagonal elements of T are positive. Define

(10) $C = [\mathrm{diag}(T)]^{-\frac{1}{2}} T [\mathrm{diag}(T)]^{-\frac{1}{2}}$; and

(11) $W = UD[\mathrm{diag}(T)]^{\frac{1}{2}}$.

Thus,

(12) $A = WCW'$,

and the $k + 1$ -order similarities matrix is given by

$$R^{(k+1)} = [\mathrm{diag}(\Lambda)]^{-\frac{1}{2}} A [\mathrm{diag}(A)]^{-\frac{1}{2}} \quad . \tag{13}$$

Now C is positive semidefinite, since (4) and (10) imply

$$C = [\mathrm{diag}(T)]^{-\frac{1}{2}} U'MU [\mathrm{diag}(T)]^{-\frac{1}{2}} \quad , \tag{14}$$

where M is idempotent, therefore positive semidefinite, and, by the "law of inertia," the transformations from M to C preserve definiteness. Furthermore, it is clear from (10) that each diagonal element of C is unity. Thus, the necessary and sufficient conditions for C to be a matrix of cosines between pairs of vectors in Euclidean n -space are satisfied.

Therefore, (12) is the familiar expression for the inner-products matrix of a set of vectors, where the columns of W represent the coordinates of the vectors with respect to n oblique axes $(\underline{w}_1, \underline{w}_2, \ldots, \underline{w}_n)$, and c_{ij} is the cosine of the angle between $\underline{w}_i$ and $\underline{w}_j$.

In effect, then, the IICA method performs a transformation of the vectors (or points) that represent the judged stimuli. The nature of this transformation is as follows: The vectors at "time k " had coordinates UD with respect to n orthogonal axes. The vectors at "time k + 1 " have coordinates $[\mathrm{diag}(A)]^{-\frac{1}{2}}W$ with respect to n oblique axes. The cosines between pairs of oblique axes are the elements of C .

The effect of this transformation can be seen more clearly by noting

$$\underline{w}_i = \underline{u}_i d_{ii} \sqrt{1 - b_i^2/n} \quad , \text{ and} \tag{15}$$

$$c_{ij} = -b_i b_j / \sqrt{(n - b_i^2)(n - b_j^2)} \quad , \qquad i \neq j \quad . \tag{16}$$

In equation (15), since $b_i = \underline{u}_i'\underline{1}$, the expression under the radical attains its maximum precisely when the sum of the elements of $\underline{u}_i$ is 0 . Now $d_{ii} = \underline{u}_i' R \underline{u}_i$, which is the squared length of the vector $\underline{u}_i$. (If R were the dispersion matrix of a set of n random variables, then this quantity would be the variance of the particular linear combination of those variables represented by $\underline{u}_i$.) Thus, in order for the elements of $\underline{w}_i$ to be "large," the length of $\underline{u}_i$ must be "large," and the sum of the elements of $\underline{u}_i$, i.e., the sum of the projections of the vectors representing the judged stimuli <u>on</u> $\underline{u}_i$, must be "small."

Consider equation (16): c_{ij} is indeterminate $0/0$ if either $b_i^2 = n$ or $b_j^2 = n$,[1] but this will seldom occur with real data. As b_i <u>or</u> b_j approaches 0 , so does c_{ij} . If $|b_i|$ <u>and</u> $|b_j|$ are large, and if they have the same sign, then c_{ij} becomes large and negative; if they have opposite signs, then c_{ij} becomes large and positive.

Intuitively, equations (15) and (16) represent two "processes" which are being "applied" simultaneously to the vectors which represent the judged stimuli in IICA. Equation (15) states that bipolar axes are "lengthened" while nonbipolar axes are "shrunk." Equation (16) states that bipolar axes tend to remain orthogonal to one another, while nonbipolar axes are rotated toward or away from one another so that they tend to "collapse" into bipolar axes.

This is how IICA converges toward a single bipolar axis, as illustrated in Clark and McQuitty (1970). In the early iterations the stimulus-vectors are transformed toward bipolarity. The greater the initial departure from bipolarity, the greater the "correction factors." As bipolarity is attained in the

---

[1]Note that, if $b_j = n$, then $b_k = 0$, for all $k \neq j$.

later iterations, the IICA process becomes similar to the well-known power-method (Anderson, 1958) for extracting the largest latent root and corresponding latent vector of a real symmetric matrix. The "process" represented by (16) becomes negligible in the later iterations.

II. Alternative Methods

The fact that IICA transforms the stimulus-vectors themselves, rather than providing a solution in terms of the original configuration, seems, a priori, to be a drawback. Are the clusters revealed by the method prominent in the data, or are they "weak"--perhaps even artifacts of the method itself? Besides such theoretical questions, there are practical problems with IICA, as shown in Section IV below, which can be avoided by the alternative techniques suggested here. Moreover, these alternative techniques allow the unified treatment of dimensional and "typal" structures, as originally suggested by McQuitty (1967).

Two methods will be proposed. Neither of these methods requires additional assumptions about the initial data matrix. Both involve extracting a single bipolar factor in such a way as to display the major clusters, and both can be applied recursively to yield hierarchical solutions. Neither makes any transformation of the original stimulus configuration.

Method A

Ignoring equation (16), and concentrating on equation (15), we seek a vector $\underline{b}$, such that $\underline{b}'R\underline{b}$ is "large" and $|\underline{b}'\underline{1}|$ is "small." (Note that this $\underline{b}$ is not the $\underline{b}$ of Section I.) Proceeding rather directly, we seek to maximize $\underline{b}'R\underline{b}$ under the constraints $\underline{b}'\underline{b} = 1$ and $\underline{b}'\underline{1} = 0$. Introducing Lagrange multipliers γ_1 , γ_2 , we write

(17) $$F(\underline{b},\gamma_1,\gamma_2) = \underline{b}'R\underline{b} - \gamma_1(\underline{b}'\underline{b} - 1) - \gamma_2(\underline{b}'\underline{1}) \quad .$$

Differentiating F with respect to $\underline{b}$, γ_1 , and γ_2 , and setting the derivative equal to 0 , yields

(18) $$2R\underline{b} - 2\gamma_1\underline{b} - \gamma_2\underline{1} = \underline{0} \quad ;$$

(19) $$\underline{b}'\underline{b} - 1 = 0 \quad ; \text{ and}$$

(20) $$\underline{b}'\underline{1} = 0 \quad .$$

Premultiplying (18) by $\underline{1}'$ gives

(21) $$\gamma_2 = 2\underline{1}'R\underline{b}/n \quad .$$

Premultiplying (18) by $\underline{b}'$ gives

(22) $$\gamma_1 = \underline{b}'R\underline{b} \quad .$$

Substituting for γ_2 in (18) gives

(23) $$(I - \underline{1}\underline{1}'/n)R\underline{b} = \gamma_1\underline{b} \quad ,$$

or, following our previous definition of M ,

(24) $$MR\underline{b} = \gamma_1\underline{b} \quad .$$

But $\underline{b}'\underline{1} = 0$ by (20), and it is easy to show that $\underline{b}'\underline{1} = 0$ iff $M\underline{b} = \underline{b}$. Therefore,[2] (24) is equivalent to

(25) $$MRM\underline{b} = \gamma_1\underline{b} \quad .$$

From (19), (22), and (25), it follows that the desired solution, $\underline{b}^*$, is the normalized latent vector of MRM corresponding to the largest latent root. But what is MRM ? It is simply the scalar-products matrix of the stimulus-

---

[2] This step, which shortens the derivation of the solution by about one page, was suggested by Dr. Walter Kristof.

vectors with respect to a coordinate system with origin at their centroid (see Torgerson, 1958, p. 258, equation 14).

The vector <u>b</u>* represents the most salient linear contrast between subsets of the stimuli. The stimuli can be ordered with respect to the corresponding elements of <u>b</u>* . Two criteria which could then be used to partition the stimuli into subsets (clusters) are

1. <u>Weak Criterion</u>: Examine the n - 1 differences between adjacent elements of <u>b</u>* (having arranged these elements in order of magnitude), find the largest such difference, and make the split between the corresponding stimuli. This should suffice when the clusters are fairly distinct.

2. <u>Strong Criterion</u>: Consider each of the n - 1 possible splits between pairs of stimuli which are adjacent with respect to <u>b</u>* . For each possible split, the original similarities matrix R can be partitioned into similarities <u>within</u> clusters and similarities <u>between</u> clusters. Thus, for each split, a quantity can be computed which reflects the adequacy of the partition. It is naturally desirable to have <u>large</u> similarities within clusters and <u>small</u> similarities between clusters. Therefore, one formula which could be used would be formally identical to the formula for the <u>alpha</u> level by a median test. Choose the split which minimizes the "alpha" for the appropriate "one-tailed test." Of course, it is not suggested that the minimum "alpha" reflects the statistical significance of the clustering. It simply provides an intuitively appealing objective function which is sensitive to cluster size as well as to differences in magnitude of similarities.

Certainly other methods of partitioning could be suggested. Examination of <u>b</u>* itself, however, will often indicate the presence or absence of clear clusters, and should provide a check on the adequacy of any method of partitioning.

Basically, what is being suggested is to use a one-dimensional scaling solution as a heuristic to reduce the number of possible partitions to be considered.

Method B

The following method constrains $|\underline{b}'\underline{1}|$ to be "small"--but not necessarily 0. Since we want to make $\underline{b}'R\underline{b}$ "large" and $|\underline{b}'\underline{1}|$ "small" at the same time, a natural function to maximize is simply $\underline{b}'R\underline{b} - (\underline{b}'\underline{1})^2$, again under the constraint $\underline{b}'\underline{b} = 1$. But since the largest possible value of $\underline{b}'R\underline{b}$ is the largest latent root of R, which can be no larger than $tr(R)$, and the largest possible value of $(\underline{b}'\underline{1})^2$ is n, a more "balanced" objective function is

$$(26) \qquad F(\underline{b}) = \underline{b}'R\underline{b} - [tr(R)/n]\,(\underline{b}'\underline{1})$$

$$= \underline{b}'\{R - [tr(R)/n]\underline{1}\underline{1}'\}\underline{b} \quad ,$$

or, letting $R^* = R - [tr(R)/n]\underline{1}\underline{1}'$,

$$(26') \qquad \underline{b}'R^*\underline{b} \ , \ \underline{b}'\underline{b} = 1 \quad .$$

The desired solution is simply the normalized latent vector corresponding to the largest latent root of R^*. This solution is invariant under transformations of the form $r_{ij} \rightarrow ar_{ij} + c$, while the Method A solution is invariant only under transformations of the form $r_{ij} \rightarrow ar_{ij}$.

The extraction of two or more vectors of R^* or MRM may be useful when cross-classification, rather than general hierarchical clustering, is desired. Cross-classification is a special case of general hierarchical clustering, since each of two subsets is partitioned with respect to the same dimension or feature, whereas in general two different subsets would be partitioned with respect to different dimensions. The variables can easily be plotted with respect to an orthogonal coordinate system that displays the major clusters.

III. Relations Among Linear Contrasts Based on Different Subsets of the Stimuli

Suppose that a set of n stimuli has been partitioned into two subsets, S_1 and S_2, according to a "contrast component" $\underline{b}$ derived by Method A or Method B. Now S_1, which contains, say, n_1 stimuli, can be further subdivided according to a contrast component $\underline{b}_1$.

An important question, first raised by McQuitty (1967, cf., his Figure 1), is, "How are $\underline{b}$ and $\underline{b}_1$ related?" In particular, are $\underline{b}$ and $\underline{b}_1$ orthogonal or not? At first this question seems meaningless, since $\underline{b}$ has n elements, while $\underline{b}_1$ has $n_1 < n$ elements. How can a scalar product be computed between two vectors that have different numbers of elements?

The following solution to this problem takes advantage of the fact that the stimuli have been partitioned with respect to identifiable underlying linear contrasts.

Let $R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$ be the partitioned similarities matrix. The vector $\underline{b}$ has been computed using R, while the vector $\underline{b}_1$ is based only on R_{11}. A vector $\underline{d}$, of n elements, can be constructed such that the first n_1 elements of $\underline{d}$ are proportional to the elements of $\underline{b}_1$, and $\underline{d}'R\underline{d}$ is maximized under the further constraint $\underline{d}'\underline{d} = 1$.

Let $\underline{d}' = [\underline{d}_1'\underline{d}_2']$, where $\underline{d}_1 = k\underline{b}_1$ for some unknown scalar k. The problem is to find $\underline{d}_2$ and k such that

$$(27) \qquad F(k,\underline{d}_2) = [k\underline{b}_1' \quad \underline{d}_2'] \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \begin{bmatrix} k\underline{b}_1 \\ \underline{d}_2 \end{bmatrix}$$

is maximized under the constraint $k^2 + \underline{d}_2'\underline{d}_2 = 1$. Once again applying the method of Lagrange multipliers, let

$$(28) \qquad G(k,\underline{d}_2,\lambda) = ck^2 + \underline{d}_2'R_{22}\underline{d}_2 + 2k\underline{v}'\underline{d}_2 - \lambda(k^2 + \underline{d}_2'\underline{d}_2 - 1) \quad ,$$

where $c = \underline{b}_1'R_{11}\underline{b}_1$ and $\underline{v} = R_{21}\underline{b}_1$.

Differentiating G with respect to k , $\underline{d}_2$, and λ , and setting the derivatives equal to 0 yields

$$(29) \qquad \underline{g}_1 = 2R_{22}\underline{d}_2 + 2k\underline{v} - 2\lambda\underline{d}_2 = \underline{0} \quad ;$$

$$(30) \qquad g_2 = 2ck + 2\underline{v}'\underline{d}_2 - 2\lambda k = 0 \quad ;$$

$$(31) \qquad g_3 = 1 - k^2 - \underline{d}_2'\underline{d}_2 = 0 \quad .$$

It may be safely assumed that $k \neq 0$. Premultiplying (29) by $\underline{d}_2'$ and dividing by 2 yields

$$(32) \qquad \underline{d}_2'R_{22}\underline{d}_2 + k\underline{d}_2'\underline{v} - \lambda\underline{d}_2'\underline{d}_2 = 0 \quad .$$

Multiplying (30) by k and dividing by 2 yields

$$(33) \qquad ck^2 + k\underline{v}'\underline{d}_2 - \lambda k^2 = 0 \quad .$$

Adding (32) and (33) yields

$$(34) \qquad ck^2 + \underline{d}_2'R_{22}\underline{d}_2 + 2k\underline{v}'\underline{d}_2 - \lambda(k^2 + \underline{d}_2'\underline{d}_2) = 0 \quad .$$

Equations (34) and (31) imply

$$(35) \qquad \lambda = ck^2 + \underline{d}_2'R_{22}\underline{d}_2 + 2k\underline{v}'\underline{d}_2 = F(k,\underline{d}_2) \quad .$$

Now (29) and (30) can be rewritten more conveniently as

$$(36) \qquad \begin{bmatrix} R_{22} & \underline{v} \\ \underline{v}' & c \end{bmatrix} \begin{bmatrix} \underline{d}_2 \\ k \end{bmatrix} = \lambda \begin{bmatrix} \underline{d}_2 \\ k \end{bmatrix} \quad .$$

From (31), (35), and (36), it follows that $\begin{bmatrix} \underline{d}_2 \\ k \end{bmatrix}$ is the normalized latent vector corresponding to the largest latent root of $\begin{bmatrix} R_{22} & \underline{v} \\ \underline{v}' & c \end{bmatrix}$.

When the values of $\underline{d}_2$ and k have been determined to the desired degree of accuracy, then the solution to the original problem is

$$(37) \qquad \underline{d}' = [\underline{d}_1' \quad \underline{d}_2'] \text{ , where } \underline{d}_1 = k\underline{b}_1 \quad .$$

Now $\underline{d}$, unlike $\underline{b}_1$, has the same number of elements as $\underline{b}$. The cosine of the angle between $\underline{b}$ and $\underline{d}$ is simply $\underline{b}'R\underline{d}/[(\underline{b}'R\underline{b})(\underline{d}'R\underline{d})]^{\frac{1}{2}}$.

IV. Example

The similarities matrix for this example (Table 1) contains phi-coefficients between pairs of subjects, based on 90 binary responses. The data were obtained in a study of reading comprehension. There were three "treatments," indicated in the tables and figures by X , Y , and Z . One of the basic hypotheses of the study was that subjects within a treatment group would be relatively homogeneous in terms of their response patterns, and that the groups would be distinct. In other words, there was an a priori three-cluster hypothesis with respect to the subject-space.

Insert Table 1 about here

Table 2 gives the coordinates of the subjects on four contrast components, namely the first and second Method A and the first and second Method B components (A1, A2, B1, and B2, respectively). In each case, the subjects have been ordered with respect to their coordinates. Differences between successive pairs of coordinates are given, and partitions have been made according to the "weak criterion" suggested in Section II.

The right-hand column of Table 2 indicates the first partition of the subjects according to IICA. Note that this split is the only one which fails to conform to the initial hypothesis. Furthermore, this misclassification by IICA is rather "robust," persisting even when the similarities are converted from phi-coefficients to rank scores.

Insert Table 2 about here

The extension of A2 and B2, from a 19-subject subspace to the full 28-subject space, is outlined in Table 3. k was found to be 0.2675 for Method A and 0.2705 for Method B. The cosine of the angle between the first and extended second components was found to be -.7753 for Method A and 0.7860 for Method B.

Insert Table 3 about here

Figure 1 shows a plot of the subjects in the plane determined by the first and extended second components according to Method A. Figure 2 shows a similar plot for the Method B components. The appropriate columns of Table 3 have been scaled according to $\underline{b}'R\underline{b}$ or $\underline{d}'R\underline{d}$.

Insert Figures 1 and 2 about here

Figures 3 and 4 show plots of two-dimensional orthogonal solutions obtained by Methods A and B, respectively. The three-cluster structure is apparent.

Insert Figures 3 and 4 about here

V. Summary and Conclusions

McQuitty's (1967) technique of hierarchical cluster analysis--Iterative Intercolumnar Correlational Analysis--has been discussed in terms of matrix algebra and geometry. Under this interpretation, it has been shown that IICA achieves a solution by transforming the stimulus-vectors themselves toward a bipolar, one-factor structure. Two alternative methods were suggested for extracting a single bipolar factor directly from the initial similarities matrix. Extension of linear contrasts from smaller to larger subspaces was also discussed. The major features of the new methods were illustrated in the analysis of some data from a reading comprehension study.

References

Anderson, T. W. An introduction to multivariate statistical analysis. New York: Wiley, 1958. Pp. 301-306.

Clark, J. A., & McQuitty, L. L. Some problems and elaborations of iterative, intercolumnar correlational analysis. Educational and Psychological Measurement, 1970, 30, 773-784.

Coles, G. J., & Stone, L. A. A new methodological revision of Ekman's "content" model of multidimensional similarity analysis. Multivariate Behavioral Research, 1972, 7, 85-107.

Johnson, S. C. Metric clustering. Unpublished paper, 1968.

Jöreskog, K. G., & van Thillo, M. New rapid algorithms for factor analysis by unweighted least squares, generalized least squares, and maximum likelihood. Research Memorandum 71-5. Princeton, N. J.: Educational Testing Service, 1971.

McQuitty, L. L. A novel application of the coefficient of correlation in the isolation of both typal and dimensional constructs. Educational and Psychological Measurement, 1967, 27, 591-599.

McQuitty, L. L. A comparative study of some selected methods of pattern analysis. Educational and Psychological Measurement, 1971, 31, 607-626.

McQuitty, L. L., & Clark, J. A. Clusters from iterative, intercolumnar correlational analysis. Educational and Psychological Measurement, 1968, 28, 211-238.

McQuitty, L. L., Abeles, N., & Clark, J. A. A study of the reliability of intra-individual personality structure by iterative intercolumnar correlational analysis. Multivariate Behavioral Research, 1970, 5, 159-176.

Torgerson, W. S. Theory and methods of scaling. New York: Wiley, 1958.

Table Headings

Table 1. Similarities matrix. *Phi*-coefficients between pairs of subjects, based on 90 binary responses. The partitioning indicates the *a priori* three-cluster hypothesis. (Decimal points are suppressed.)

Table 2. Coordinates of subjects on contrast components. Subjects have been ordered with respect to their coordinates, and differences between successive pairs of coordinates are given in the column to the right of the coordinates themselves. Partitions according to the weak criterion are given. The right-hand column gives the first partition according to IICA. (Decimal points are suppressed.)

Table 3. Extension of second Method A and Method B contrast components. Blanks in columns A2 and B2 are filled by the extension $\underline{d}_2$. Other elements in columns A2E and B2E are equal to k times the corresponding elements of columns A2 and B2, respectively.

Table 1

Similarities Matrix

	X1	X2	X3	X4	X5	X6	X7	X8	X9	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9
X1	100	54	58	49	59	56	44	47	82	9	39	46	5	36	-1	33	20	22
X2	54	100	76	47	76	68	66	61	57	18	13	16	23	19	-7	27	22	20
X3	58	76	100	61	76	68	62	63	47	19	17	25	28	15	0	22	31	20
X4	49	47	61	100	47	34	52	63	48	17	22	29	13	35	14	30	15	23
X5	59	76	76	47	100	87	76	66	62	13	13	21	27	9	-7	17	22	11
X6	56	68	68	34	87	100	65	57	59	2	5	13	16	4	-11	5	9	-2
X7	44	66	62	52	76	65	100	62	47	16	2	11	36	8	9	6	18	15
X8	47	61	63	63	66	57	62	100	50	25	10	14	15	32	0	21	19	30
X9	82	57	47	48	62	59	47	50	100	4	30	37	8	29	7	28	11	12
Y1	9	18	19	17	13	2	16	25	4	100	51	54	54	48	51	61	74	73
Y2	39	13	17	22	13	5	2	10	30	51	100	88	42	54	39	65	64	71
Y3	46	16	25	29	21	13	11	14	37	54	88	100	36	42	32	59	67	60
Y4	5	23	28	13	27	16	36	15	8	54	42	36	100	24	60	38	61	41
Y5	36	19	15	35	9	4	8	32	29	48	54	42	24	100	32	72	35	65
Y6	-1	-7	0	14	-7	-11	9	0	7	51	39	32	60	32	100	36	37	39
Y7	33	27	22	30	17	5	6	21	28	61	65	59	38	72	36	100	54	71
Y8	20	22	31	15	22	9	18	19	11	74	64	67	61	35	37	54	100	65
Y9	22	20	20	23	11	-2	15	30	12	73	71	60	41	65	39	71	65	100
Z1	-1	8	0	3	2	9	-6	19	0	25	7	-7	-1	20	-3	15	17	20
Z2	30	17	12	19	7	13	-3	22	25	24	29	18	5	57	12	40	11	33
Z3	20	14	5	9	0	8	-2	17	15	25	23	15	-3	48	9	29	18	28
Z4	22	11	2	13	4	5	-4	14	24	15	11	10	-19	30	-8	18	14	18
Z5	22	22	1	9	2	8	-5	19	23	20	19	19	-20	36	-7	18	15	17
Z6	5	-12	-13	-3	-12	-10	-20	-2	7	-1	7	6	-14	27	3	15	-9	8
Z7	10	15	7	16	9	9	3	22	24	18	5	-2	-12	28	-1	15	10	20
Z8	7	21	4	1	5	8	-6	14	9	18	11	-5	-5	26	-16	23	19	15
Z9	2	12	1	9	2	-2	-5	19	3	30	19	-1	0	36	-7	28	25	27
Z10	15	17	6	-5	7	13	2	5	16	24	21	12	-4	40	-1	22	18	21

Table 1 (Continued)

	Z1	Z2	Z3	Z4	Z5	Z6	Z7	Z8	Z9	Z10
X1	-1	30	20	22	22	5	10	7	2	15
X2	8	17	14	11	22	-12	15	21	12	17
X3	0	12	5	2	1	-13	7	4	1	6
X4	3	19	9	13	9	-3	16	1	9	-5
X5	2	7	0	4	2	-12	9	5	2	7
X6	9	13	8	5	8	-10	9	8	-2	13
X7	-6	-3	-2	-4	-5	-20	3	-6	-5	2
X8	19	22	17	14	19	-2	22	14	19	5
X9	0	25	15	24	23	7	24	9	3	16
Y1	25	24	25	15	20	-1	18	18	30	24
Y2	7	29	23	11	19	7	5	11	19	21
Y3	-7	18	15	10	19	6	-2	-5	-1	12
Y4	-1	5	-3	-19	-20	-14	-12	-5	0	-4
Y5	20	57	48	30	36	27	28	26	36	40
Y6	-3	12	9	-8	-7	3	-1	-16	-7	-1
Y7	15	40	29	18	18	15	15	23	28	22
Y8	17	11	18	14	15	-9	10	19	25	18
Y9	20	33	28	18	17	8	20	15	27	21
Z1	100	38	52	49	45	19	58	71	59	40
Z2	38	100	33	29	26	25	36	36	26	29
Z3	52	33	100	45	65	23	47	55	50	58
Z4	49	29	45	100	48	12	61	31	33	42
Z5	45	26	65	48	100	31	41	57	58	71
Z6	19	25	23	12	31	100	24	18	4	27
Z7	58	36	47	61	41	24	100	55	41	48
Z8	71	36	55	31	57	18	55	100	73	50
Z9	59	26	50	33	58	4	41	73	100	52
Z10	40	29	58	42	71	27	48	50	52	100

Table 2

Coordinates of Subjects on Contrast Components

S#	A1	diff	S#	A2	diff	S#	B1	diff	S#	B2	diff	IICA
Z1	-239	005	Y4	-323	035	X5	-312	007	Y4	-323	034	Z6
Z8	-234	000	Y3	-288	013	X7	-305	016	Y3	-289	014	Z2
Z9	-234	010	Y6	-274	018	X3	-288	017	Y6	-274	017	Z9
Z3	-223	001	Y2	-256	006	X6	-271	033	Y2	-257	006	Z8
Z5	-222	008	Y8	-250	031	X2	-238	041	Y8	-251	031	Z10
Z10	-214	026	Y9	-219	001	X8	-197	004	Y9	-221	001	Z5
Z7	-188	012	Y1	-218	027	X4	-193	000	Y1	-220	027	Z3
Z4	-176	005	Y7	-191	126	X1	-193	001	Y7	-193	125	Z7
Z6	-171	055	Y5	-066	134	X9	-193	081	Y5	-067	135	Z4
Z2	-116	010	Z2	069	056	Y4	-112	062	Z2	068	057	Z1
Y5	-106	047	Z6	124	077	Y3	-050	048	Z6	124	076	Y7
Y1	-059	001	Z3	202	011	Y8	-002	012	Z3	200	011	Y5
Y9	-058	012	Z9	213	003	Y6	009	013	Z9	211	003	Y9
Y7	-047	020	Z10	215	002	Y2	023	019	Z10	214	003	Y1
Y2	-027	017	Z4	217	037	Y7	042	012	Z4	217	036	Y6
Y6	-010	009	Z5	254	003	Y9	054	002	Z5	253	003	Y4
Y8	-002	048	Z7	257	002	Y1	055	047	Z7	256	002	Y8
Y3	046	064	Z1	259	017	Y5	102	012	Z1	258	017	Y3
Y4	110	080	Z8	276		Z2	113	058	Z8	275		Y2
X9	190	000				Z6	172	003				X8
X1	190	001				Z4	175	012				X4
X4	191	003				Z7	186	025				X7
X8	194	041				Z10	212	008				X3
X2	235	034				Z5	220	001				X6
X6	270	016				Z3	221	012				X5
X3	286	018				Z9	231	001				X2
X7	303	006				Z8	232	006				X1
X5	309					Z1	238					X9

Table 3

Extension of Second Method A and Method B Contrast Components

S#	A1	A2	A2E	B1	B2	B2E
X1	190	----	-302	-193	----	-302
X2	235	----	-332	-238	----	-332
X3	286	----	-349	-288	----	-349
X4	191	----	-279	-193	----	-279
X5	310	----	-365	-312	----	-364
X6	270	----	-322	-271	----	-322
X7	303	----	-329	-305	----	-328
X8	194	----	-307	-197	----	-306
X9	190	----	-298	-193	----	-298
Y1	-060	-219	-059	055	-220	-060
Y2	-027	-256	-068	023	-257	-070
Y3	046	-288	-077	-050	-289	-078
Y4	110	-323	-086	-112	-323	-087
Y5	-106	-066	-018	102	-067	-018
Y6	-010	-274	-073	009	-274	-074
Y7	-047	-191	-051	042	-193	-052
Y8	-002	-250	-067	-002	-251	-068
Y9	-058	-219	-059	054	-221	-060
Z1	-239	259	069	238	258	070
Z2	-116	069	018	113	068	018
Z3	-223	201	054	221	200	054
Z4	-176	217	058	175	217	059
Z5	-222	254	068	220	253	068
Z6	-171	124	033	172	124	034
Z7	-188	257	069	186	256	069
Z8	-234	276	074	232	275	074
Z9	-234	213	057	231	211	057
Z10	-214	215	058	212	214	057

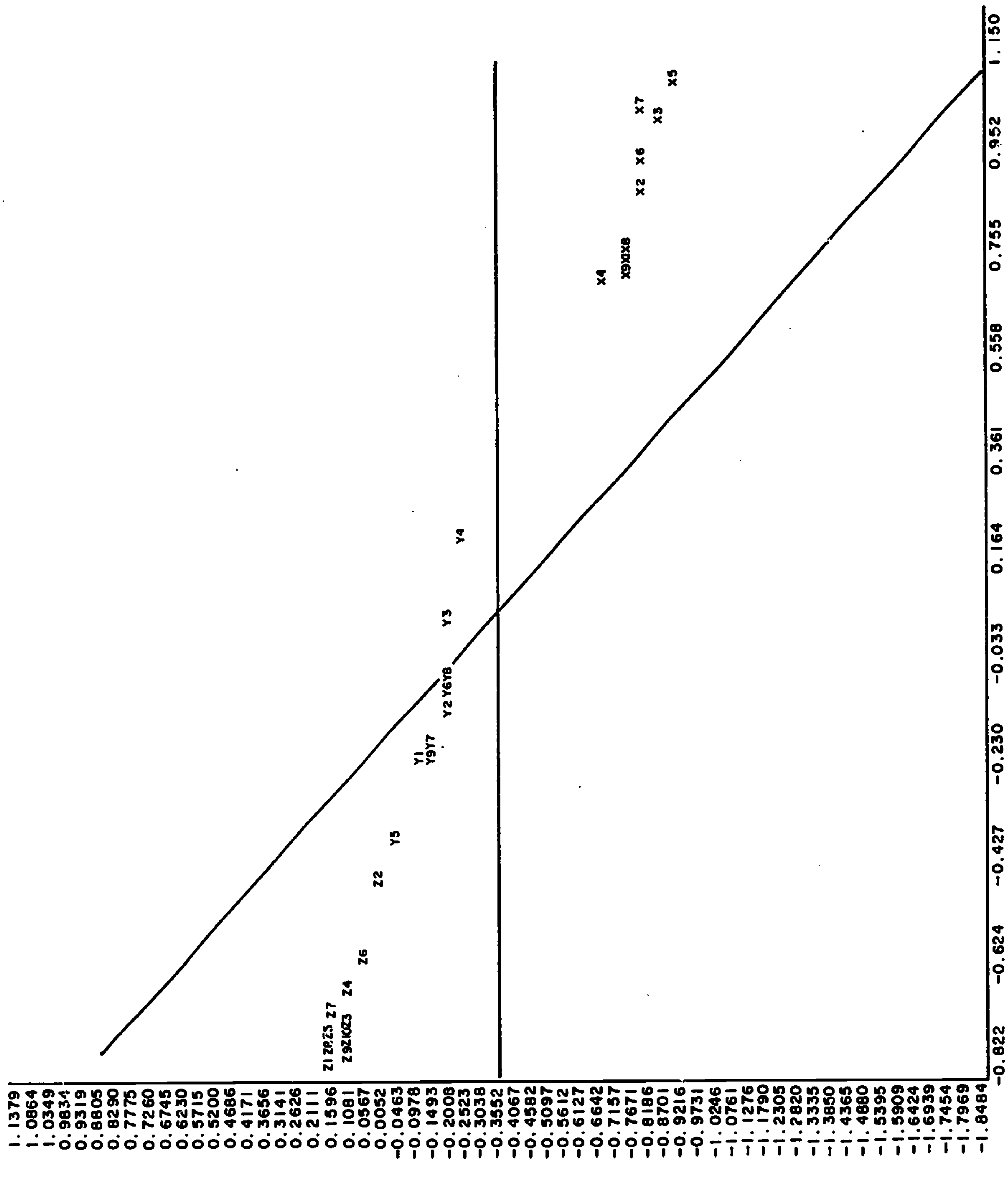
Figure Captions

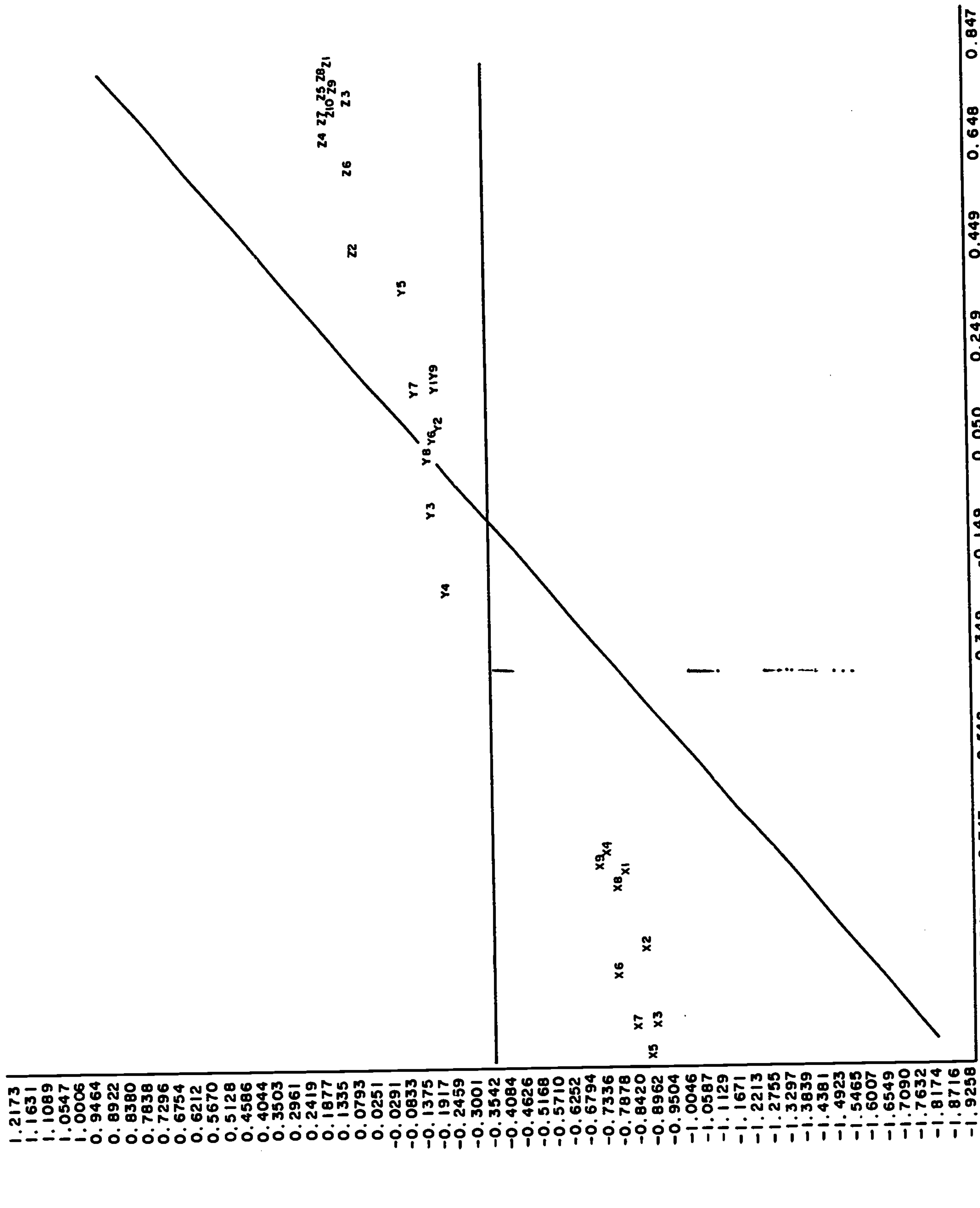
Fig. 1. Plot of columns A1 (b) and A2E (d) from Table 3. The coordinates have been multiplied by the square root of "variance-accounted-for." The cosine of the angle between the axes is -.76.

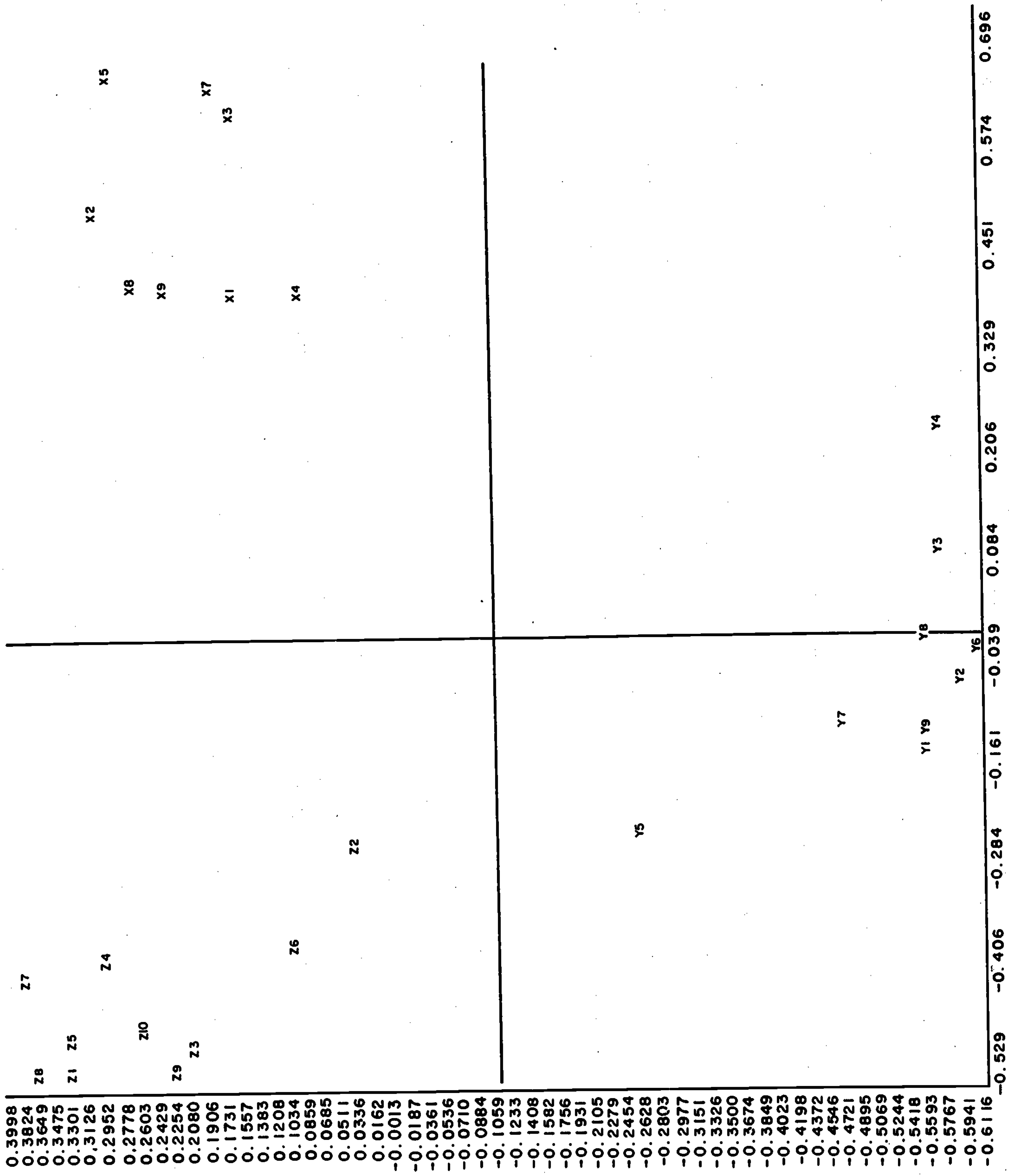
Fig. 2. Plot of columns B1 (b) and B2E (d) from Table 3. The coordinates have been multiplied by the square root of "variance-accounted-for." The cosine of the angle between the axes is 0.79.

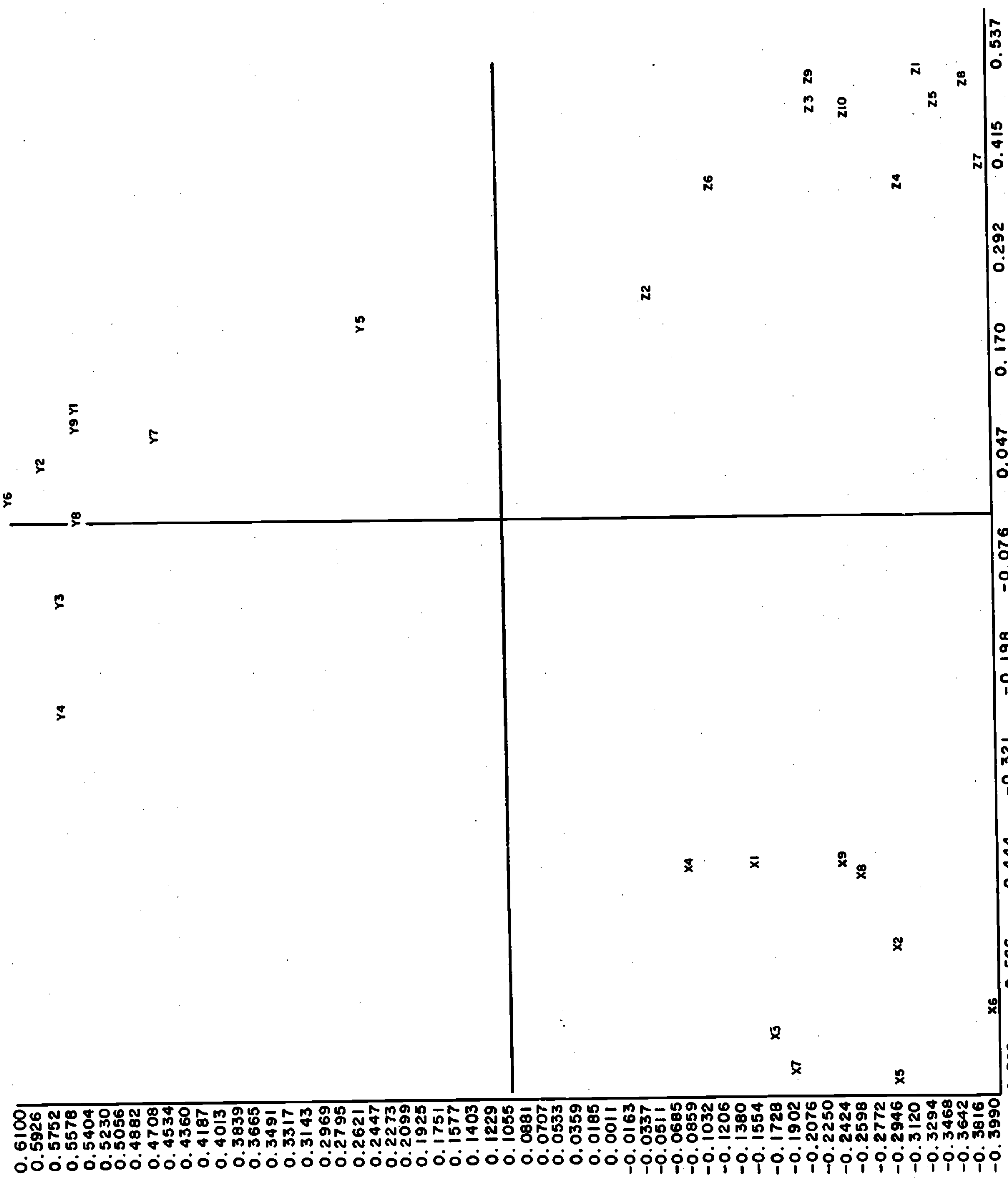
Fig. 3. Two-dimensional orthogonal solution, Method A.

Fig. 4. Two-dimensional orthogonal solution, Method B.


1.1379
1.0864
1.0349
0.9834
0.9319
0.8805
0.8290
0.7775
0.7260
0.6745
0.6230
0.5715
0.5200
0.4686
0.4171
0.3656
0.3141
0.2626
0.2111
0.1596
0.1081
0.0567
0.0052
-0.0463
-0.0978
-0.1493
-0.2008
-0.2523
-0.3038
-0.3552
-0.4067
-0.4582
-0.5097
-0.5612
-0.6127
-0.6642
-0.7157
-0.7671
-0.8186
-0.8701
-0.9216
-0.9731
-1.0246
-1.0761
-1.1276
-1.1790
-1.2305
-1.2820
-1.3335
-1.3850
-1.4365
-1.4880
-1.5395
-1.5909
-1.6424
-1.6939
-1.7454
-1.7969
-1.8484
-0.822
-0.624
-0.427
-0.230
-0.033
0.164
0.361
0.558
0.755
0.952
1.150
Z4
Z6
Z2
Y5
Y1
Y9Y7
Y2 Y6Y8
Y3
Y4
X4
X9X1X8
X2
X6
X7
X3
X5


0.847
0.648
0.449
0.249
0.050
-0.149
-0.349
-0.548
-0.747
-0.947
-1.146
1.2173
1.1631
1.1089
1.0547
1.0006
0.9464
0.8922
0.8380
0.7838
0.7296
0.6754
0.6212
0.5670
0.5128
0.4586
0.4044
0.3503
0.2961
0.2419
0.1877
0.1335
0.0793
0.0251
-0.0291
-0.0833
-0.1375
-0.1917
-0.2459
-0.3001
-0.3542
-0.4084
-0.4626
-0.5168
-0.5710
-0.6252
-0.6794
-0.7336
-0.7878
-0.8420
-0.8962
-0.9504
-1.0046
-1.0587
-1.1129
-1.1671
-1.2213
-1.2755
-1.3297
-1.3839
-1.4381
-1.4923
-1.5465
-1.6007
-1.6549
-1.7090
-1.7632
-1.8174
-1.8716
-1.9258
Z4 Z7 Z5 Z8 Z1
Z10 Z9
Z3
Z6
Z2
Y5
Y7
Y1Y9
Y8 Y6 Y2
Y3
Y4
X9 X4
X8 X1
X2
X6
X7
X3
X5


X1 X2 X3 X4 X5 X6 X7 X8 X9
Y1 Y2 Y3 Y4 Y5 Y6 Y7 Y8 Y9
Z1 Z2 Z3 Z4 Z5 Z6 Z7 Z8 Z9 Z10
0.696 0.574 0.451 0.329 0.206 0.084 -0.039 -0.161 -0.284 -0.406 -0.529
0.3998 0.3824 0.3649 0.3475 0.3301 0.3126 0.2952 0.2778 0.2603 0.2429 0.2254 0.2080 0.1906 0.1731 0.1557 0.1383 0.1208 0.1034 0.0859 0.0685 0.0511 0.0336 0.0162 -0.0013 -0.0187 -0.0361 -0.0536 -0.0710 -0.0884 -0.1059 -0.1233 -0.1408 -0.1582 -0.1756 -0.1931 -0.2105 -0.2279 -0.2454 -0.2628 -0.2803 -0.2977 -0.3151 -0.3326 -0.3500 -0.3674 -0.3849 -0.4023 -0.4198 -0.4372 -0.4546 -0.4721 -0.4895 -0.5069 -0.5244 -0.5418 -0.5593 -0.5767 -0.5941 -0.6116


0.537
0.415
0.292
0.170
0.047
-0.076
-0.198
-0.321
-0.444
-0.566
-0.689
0.6100
0.5926
0.5752
0.5578
0.5404
0.5230
0.5056
0.4882
0.4708
0.4534
0.4360
0.4187
0.4013
0.3839
0.3665
0.3491
0.3317
0.3143
0.2969
0.2795
0.2621
0.2447
0.2273
0.2099
0.1925
0.1751
0.1577
0.1403
0.1229
0.1055
0.0881
0.0707
0.0533
0.0359
0.0185
0.0011
-0.0163
-0.0337
-0.0511
-0.0685
-0.0859
-0.1032
-0.1206
-0.1380
-0.1554
-0.1728
-0.1902
-0.2076
-0.2250
-0.2424
-0.2598
-0.2772
-0.2946
-0.3120
-0.3294
-0.3468
-0.3642
-0.3816
-0.3990
Y6
Y2
Y9 Y1
Y8
Y7
Y3
Y4
Y5
Z2
Z6
Z3 Z9
Z10
Z1
Z5
Z8
Z4
Z7
X4
X1
X9
X8
X2
X6
X3
X7
X5